**Effects of nucleoid-associated proteins on bacterial chromosome structure and gene expression**

Douglas F Browning\(^1\), David C Grainger\(^2\) and Stephen JW Busby\(^3\)

Bacterial nucleoid-associated proteins play a key role in the organisation, replication, segregation, repair and expression of bacterial chromosomes. Here, we review some recent progress in our understanding of the effects of these proteins on DNA and their biological role, focusing mainly on *Escherichia coli* and its chromosome. Certain nucleoid-associated proteins also regulate transcription initiation at specific promoters, and work in concert with dedicated transcription factors to regulate gene expression in response to growth phase and environmental change. Some specific examples, involving the *E. coli* IHF and Fis proteins, that illustrate new principles, are described in detail.

**Addresses**

\(^1\) School of Immunity and Infection, University of Birmingham, Birmingham B15 2TT, UK  
\(^2\) Department of Biological Sciences, University of Warwick, Coventry CV4 7AL, UK  
\(^3\) School of Biosciences, University of Birmingham, Birmingham B15 2TT, UK

Corresponding author: Busby, Stephen JW ([s.j.w.busby@bham.ac.uk](mailto:s.j.w.busby@bham.ac.uk))

**Introduction**

The nucleoid is a region of the bacterial cell into which its chromosome is constrained. To fit into this space, the DNA has to be highly compacted and this is due to the action of supercoiling, RNA and nucleoid-associated proteins (NAPs). Bacterial chromosome folding and the factors involved in their folding began to be investigated in the 1970s. However, the first thorough and systematic investigation of NAPs came from Akira Ishihama and colleagues in 1999 [1,2], who catalogued the different NAPs of *Escherichia coli*, measured their binding affinities, investigated their preferred targets, quantified their levels, and reported dramatic changes in the levels of certain NAPs in response to changes in growth. The arrival of chromatin immunoprecipitation technology together with whole genome sequences has led to a picture of the distribution of individual NAPs across whole chromosomes [3], whilst the total protein occupancy landscape has been established using complementary experimental approaches [4,5]. Remarkably, we are still largely ignorant of precisely how bacterial chromosomes are packaged and organised, but the literature, perhaps mistakenly, places a big emphasis on the importance of NAPs [6]. Here, we present a selection of recent reports, describing effects of NAPs on bacterial chromosome structure and gene expression, and we refer the reader to the excellent review by Dillon and Dorman [7\(^*\)] for a comprehensive up-to-date picture.

**Bacterial folded chromosomes**

Of all the bacterial chromosomes, it is the *E. coli* chromosome that has been most studied. Its packaging and organisation can be probed using fluorescent proteins targeted to specific loci and Sherratt and colleagues’ recent review [8] explains how its replication and segregation are organised. Wiggins *et al.* [9] have pushed the technology to the limit and measured the distribution and fluctuations of many different locations and their juxtaposition, and they conclude that loci show remarkable precision in their positioning. Thus, although many textbooks sketch bacterial chromosomes as a more or less disordered jumble of DNA, this is far from the reality and they are likely to be organised with the same precision as other macromolecular assemblies, although with a built-in facility for change due to the mechanics of DNA replication, repair, segregation and expression. With respect to building this structure, Sexton *et al.* [10] argue for the importance of meeting points between different locations in chromosomes, and Junier *et al.* [11] describe the physical consequences of these interactions. The resulting periodicity, and the creation of different neighbourhoods within the same cell, may have consequences for gene expression and for the organisation of regulatory networks, and these are explored by Janga *et al.* [12] and Mathelier and Carbone [13].

The traditional view of the *E. coli* folded chromosome is that it consists of a series of constrained loops separated by dynamic boundaries. Boccard and his colleagues have suggested a higher order of organisation in which the chromosome is partitioned into six discrete ‘macrodomains’, and fluorescence microscopy shows that the local DNA dynamics differs in each of the macrodomains that may correspond to distinct physical structures [14]. In a recent study [15\(^*\)], Boccard’s team identified a new nucleoid-associated protein, MatP, whose binding is specific for the *ter* macrodomain that covers the replica-
tion termination region. This is an important result as it shows that the local structure of each segment of a bacterial chromosome might not be identical, though, for the moment, it is not clear how MatP binding confers distinctive properties on the ter macrodomain.

**Flexibility and variation in NAPs**

Studies of different bacterial systems in different conditions have revealed complexities that were not apparent in Ishihama and colleagues’ original ‘audit’ of *E. coli* NAPs [1,2]. For example, the observation that the high levels of Fis in rapidly growing cells are reduced to near zero as cell growth slows does not apply in anaerobic cultures of *Salmonella enterica* serovar Typhimurium [16]. Moreover, there is great variation in NAPs in different bacteria and Fis is restricted to Gamma proteobacteria [17], although AbrB appears to play the same role in *Bacillus subtilis* [18]. Takeyasu and colleagues have exploited atomic force microscopy [19] and optical tweezers [20] to observe big changes in the fibre structure of the *E. coli* chromosomes in different growth conditions and have attributed these to changes in the levels of different NAPs [17] and also transcription patterns [21,22]. However the behaviour of *E. coli* is more elaborate than many bacteria and, for example, such changes are not seen in the chromosomes of *Staphylococcus aureus* [23]. Finally the lists of different NAPs found in the literature for different bacteria may not yet be a closed book. Thus the *B. subtilis* Noc protein [24] and the *E. coli* Dan protein [25] are recent additions, and the discovery of MatP [15] hints that there may be many more locally acting NAPs to be found.

**NAPs and DNA structure**

*In vitro* studies with purified individual NAPs using both ensemble and single molecule measurements show that bending, bridging, wrapping and clustering can result, following binding to DNA [6]. Some of these modes are illustrated in Figure 1. For example, DNA bridging by H-NS has been most comprehensively by Remus Dame and his colleagues using a variety of biophysical methods [26]. A recent report suggests that different modes of H-NS action could be controlled by divalent metal ions [27]. Similarly, HU, the most universally conserved and abundant NAP, can induce DNA bends, condense DNA in a fibre, and also interact with single stranded DNA [28,29]. Combinations of NAPs have been studied systematically by Muskhelishvili and colleagues and shown to generate novel types of DNA organisation [30]. Significantly, Dps and CbpA, both of which are strongly induced in stationary phase *E. coli*, cause compaction by clustering of distal DNA loci (Figure 1b) [31].

**NAPs and global transcription**

Since most of the NAPs are high abundance and quite promiscuous with respect to DNA binding it is unsurprising that they affect transcription on a global scale. For example, two recent reports [32,33] underscore the profound influence of HU in coordinating the *E. coli* transcriptome. The importance of NAPs in global transcription is reinforced by chromatin immunoprecipitation experiments which showed that, for H-NS, IHF and Fis in *E. coli*, the majority of binding targets were located in intergenic regulatory regions [34]. Effects on global transcription not only occur due to changes in the relative levels of NAPs in different growth phases, but also arise due to NAP binding being affected by the passage of replication forks during the cell cycle. For example, a recent report [35] highlighted a possible role for SeqA, which binds to targets containing hemi-methylated GATC motifs.

Much recent literature attention has focused on H-NS, which is often referred to as the ‘genome guardian’, the ‘universal repressor’ or a ‘sentinel’ [7-36]. H-NS presents a tantalising problem, because, despite having been first reported in 1971 [37], we still lack understanding, at the molecular level, of specific repression by H-NS and how it is modulated. What is clear is that H-NS recognizes AT-rich segments of the chromosome, and then binding is in two steps, an initial binding of the H-NS dimer, followed by further co-operative binding, and this can result in molecular bridging between distal chromosome locations and concomitant transcription regulation [36]. A major breakthrough was the identification of high affinity DNA sites for H-NS that are, perhaps, responsible for step one of H-NS binding [38*].

The existence of the H-NS parologue, StpA, in both *E. coli* and *S. enterica* serovar Typhimurium has long posed a puzzle, since *stpA* knockouts have minimal effects in *E. coli*. Using biochemical approaches, John Ladbury’s group showed that H-NS and StpA could form both homodimers and heterodimers and this may well explain genome-wide chromatin immunoprecipitation data from Oshima and colleagues [40], who found that the vast majority of DNA targets in *E. coli* occupied by StpA were also occupied by H-NS. Deletion of *stpA* had a minimal effect on the distribution of H-NS, whilst deletion of *hns* greatly reduced the number of StpA-bound targets. Thus, in *E. coli*, StpA binds mainly as a heterodimer with H-NS, acting almost as a ‘reserve’. By contrast, in *S. enterica* serovar Typhimurium, StpA plays a significant role, an *stpA* deletion affects expression of 5% of the genome, and there appears to be an StpA regulon [41].

The observation that many of the A:T-rich targets for H-NS are located in genome segments that have arrived by horizontal evolution prompted the idea that H-NS is a ‘sentinel’ whose task is to suppress expression from externally acquired pieces of DNA [7*]. Support for this comes from the observation that some transmissible plasmids carry a gene encoding an H-NS homologue. Doyle
et al. [42] identified Sfh as such a homologue and showed that, without Sfh, incoming plasmids stress the host cell as they titrate out host H-NS. Hence Sfh can be viewed as a molecular stealth device that permits the plasmid to enter the host without causing disruption. A recent genome-wide analysis of H-NS and Sfh binding in *S. enterica* serovar Typhimurium [43] supports this view by showing that Sfh and H-NS recognise common target sites, and argues that the invading plasmid deploys Sfh to ‘top-up’ the level of H-NS thereby minimizing disturbance to its new host. This scenario, with H-NS as a sentinel, supposes that, if a segment of horizontally acquired DNA segment encoded something that was useful for its new host, mechanisms would evolve to ensure activation. Recent reviews from Dorman and colleagues [44,45] discuss this issue and outline the different mechanisms by which repression by H-NS and its homologues at specific targets can be lifted.

Another current focus of H-NS research concerns *Mycobacteria* where initial scans of genome sequences suggested that H-NS was absent and that a protein that resembled the GroEL chaperone functioned as a NAP [46]. However it has now been shown that the *M. tuberculosis* Rv3597 and Rv3852 genes encode proteins (Lsr2 and H-NS respectively) whose *in vitro* properties resemble H-NS from enterics. Concerning Lsr2, although it shares but 4% sequence identity with *E. coli* H-NS, it is a DNA-bridging protein that binds to AT-rich regions, and can complement some effects of *hns* mutations in *E. coli* [47,48]. Similarly, the H-NS-like product of the *M. tuberculosis* Rv3852 gene can complement *hns* mutations in *E. coli* [49]. Interestingly the *M. tuberculosis* Rv3852 gene product has been shown to interact with Holliday junctions [50], and this underscores the point that, although the majority of literature concerning H-NS and its paralogues focuses on their interaction with double
stranded DNA, there is considerable evidence for interactions with single stranded nucleic acids, including messenger RNA. For example, a recent report [51] describes interactions of H-NS with the leader sequence of the *E. coli malT* gene, and argues that H-NS activates the initiation of translation of MalT, and may interact likewise at other messengers.

**IHF: a NAP that likes to play with others**

Bacterial Integration Host Factor (IHF) is a NAP that plays many roles, binds to many DNA targets, and is known to act like a transcription factor at many gene regulatory regions in *E. coli* [34]. Its major function is to bend target DNA sharply and hence it often functions in conjunction with other NAPs and transcription factors, such as at the *fim* switch that controls phase variable of type 1 fimbriae in *E. coli* [52]. Recent experiments with the *E. coli nir* and *nrf* promoters, which both are induced by anaerobiosis together with nitrite or nitrate ions, have revealed new subtleties in IHF action. Both promoters control the expression of a nitrite reductase and expression is completely dependent on FNR, the master regulator of transcription in anaerobic conditions. At both promoters, FNR-dependent transcription activation is suppressed by upstream-bound IHF. At the *nir* promoter, this suppression is effected by the combined action of IHF (which binds at position −88) and Fis (which binds at position −142). The dependence of the promoter on nitrite or nitrate ions comes from the fact that NarL and NarP, transcription activators that are triggered by nitrite and nitrate, bind to a site that overlaps the DNA site for IHF (Figure 2a). Binding displaces IHF and this results in relief of suppression of FNR-dependent *nir* promoter activity. The unexpected subtlety is the existence of a second weaker DNA site for IHF located at position −115, and that IHF binding at this site is activatory for FNR-dependent *nir* promoter activity. Hence basal *nir* promoter activity in anaerobic conditions in the absence of nitrite or nitrate depends on the relative occupation of the two DNA sites for IHF. The base sequence of these two sites varies from one *E. coli* strain to another and this variation sets the basal activity in any strain [53].

**Figure 2**

Transcription regulation at the *E. coli nir* and *nrf* operon promoters. (a) The *nir* promoter. Transcription initiation is dependent on FNR. The binding of Fis and IHF inhibits FNR-activated transcription (−ve), whilst IHF binding to a second lower affinity site stimulates transcription (+ve). The binding of NarL (or its homologue, NarP) displaces IHF, counteracting the repression mediated by IHF and Fis, and enabling maximal FNR-dependent transcription. (b) The *nrf* promoter. Transcription initiation is dependent on FNR. The binding of IHF, Fis and NsrR to the promoter inhibits FNR-activated transcription (−ve), whilst occupancy of a second upstream IHF site stimulates transcription (+ve). The binding of NarL or NarP displaces IHF, counteracting IHF-mediated repression and increasing FNR-dependent transcription. Repression by NsrR is relieved by the presence of reactive nitrogen species. In both panels, the transcription start point (+1) is indicated by a bent arrow.
Regulation of transcription initiation by Fis. The figure shows sketches to illustrate the different ways that Fis can interact at bacterial promoters. In all panels, the transcription startpoint (+1) is indicated by a bent arrow. (a) Simple class I activation. Fis directly contacts the C-terminal domain of one of the α subunits of holo RNA polymerase (Eσ70), thereby recruiting RNA polymerase to the promoter. (b) Altering DNA conformation. Binding of Fis to the upstream of the promoter stabilizes the “breathing” of the DNA helix, transmitting this energy downstream and facilitating the ease with RNA polymerase interacts with the promoter. (c) Direct repression. The binding of Fis to the promoter region directly blocks the association of holo RNA polymerase (Eσ70). (d) Displacement of an essential activator. Fis represses transcription initiation by displacing an essential activator (A). (e) Preventing the function of an activator. Fis represses transcription initiation by interfering with the ability of an activator protein (A) to activate transcription. (f) Jamming of RNA polymerase at promoters. Binding of Fis, in unison with holo RNA polymerase containing σ70 (Eσ70), shuts down transcription by creating a repression complex. This blocks access to the promoter by RNA polymerase containing the alternative “stationary phase” σ38 factor (Eσ38).

A similar situation is found at the E. coli nrf promoter, where IHF bound at position −54 represses FNR-dependent activation, whilst IHF bound further upstream at −127 promotes activation. Surprisingly, the repression of FNR-dependent activation by IHF at the nrf promoter is less efficient than at the nir promoter, and it has recently been shown that a second repressor, NsrR is involved [54]. Hence, FNR-dependent activation of the nrf pro-
moter is down-regulated independently by two different factors, IHF that binds at position –54, and NsrR that binds at position –63. Repression by IHF is lifted by NarL or NarP binding in response to nitrate or nitrite, whilst repression by NsrR is lifted when its DNA binding activity is prevented when it interacts with reactive nitrogen species generated as by-products of nitrate or nitrite metabolism (Figure 2b). A distinctive feature of the E. coli nrf promoter is that binding of Fis to a target at position –15 overrides all the other inputs from FNR, IHF, NsrR, NarL and NarP, and thus, unusually, this promoter is subject to three different methods of repression (Figure 2b).

**Fis: the busy-body of the cell**

The factor for inversion stimulation (Fis) is another DNA-bending NAP that plays many roles and binds at hundreds of DNA targets in E. coli [34,55]. It is especially important because it reaches very high levels in rapidly growing cells but it is almost absent in certain conditions of slow growth. The first crystal structure of Fis bound to a DNA target has recently been reported [56]. Detailed studies of Fis action at individual regulatory regions have revealed at least six different mechanisms by which it affects gene expression (summarised in Figure 3). At some promoters, Fis activates transcription initiation either by directly interacting with RNA polymerase thereby recruiting it to the promoter, or by inducing an activatory conformational change in the promoter [57]. However a major task for Fis appears to be to shut off the expression of inessential gene products during rapid growth. This can be done by direct binding to essential promoter elements as at the E. coli nrf promoter (Figure 2b), or by displacing an essential activator [58] or preventing the function of an activator [53]. Another mechanism of repression employed by Fis is RNA polymerase jamming. Hence at the E. coli dps promoter, that controls expression of the major stationary phase NAP, Dps, Fis combines with RNA polymerase containing the house keeping sigma 70 factor to form a repressive ternary complex [59].

**Conclusions and perspectives**

Although our understanding of NAPs and their effects has greatly advanced over the past few years, there are still many issues to be tackled. Notably, how is bacterial chromosomal DNA organised in cells and what are the consequences on replication, transcription and repair? What do the different macrodomains correspond to and how are they controlled? How far is the E. coli model applicable to other bacteria where there are a smaller number of NAP species and do other factors such as cis-acting RNA play a role? Many of these issues are dealt with in a stimulating way in a recently published edited book on bacterial chromatin [60]. Meanwhile many fresh approaches are being taken to tackle the problems, including sophisticated imaging, tomography and new analytical approaches to home in on selected segments of bacterial chromosomes [61].

**Acknowledgements**

The authors have been supported by a Wellcome Trust Programme Grant to SJWB and a Wellcome Trust Research Career Development Fellowship to DCG. We apologise that we had to omit many excellent relevant papers due to space constraints.

**References and recommended reading**

Papers of particular interest, published within the period of review, have been highlighted as:

- of special interest


This easy to read, yet comprehensive review, details the many roles that NAPs play in structuring the bacterial chromosome and regulating gene expression in different bacterial species.


Of the protein, which binds to targets specifically located in an extended region of the Escherichia coli chromosome centred on the replication terminus.


In this clever application of bioinformatics, the authors build on their previous observation that H-NS binds to DNA by recognising a specific “nucleation” sequence.


How do A-T rich plasmids sneak under the radar and avoid titrating nucleoid proteins away from the host’s chromosome? In this paper, the authors show that the “stealth” protein SfH, which functions as a plasmid encoded H-NS homologue, plays a key role. Ref. [43] describes a genome-wide study of SfH.


This paper describes experiments to show that H-NS can act as a positive effector of gene expression by interacting with a specific messenger RNA and stimulating the initiation of translation. This observation may well resolve some longstanding anomalies in the literature concerning what H-NS does and does not do.


This paper presents a thorough structural analysis, which describes how DNA conformation, as well as DNA sequence, plays a key role in permitting Fis to interact with its DNA target.


